大数据技术驱动的纪录片内容呈现技巧变化

摘　要：大数据技术的运用为纪录片的呈现提供了多种新的技术方式，丰富了纪录片的表现形式。在对大量复杂图像、技术的分析过程中运用了众包技术，在对各项信息、知识的分析过程中运用了多项数据技术方式，运用大数据优化对图像的收集、处理与运用，运用大数据中的查询引擎、数据存储与数据处理工具出具分析报道，使用便携 4K 设备进行拍摄与数据处理，优化作业流程，加强素材调色，结合纪录片的呈现方式体现技巧变化。
关键词：大数据技术；纪录片；查询引擎；4K 设备；素材调色
中图分类号：J952　文献标识码：A
文章编号：1671-0134（2019）10-052-03　DOI：10.19483/j.cnki.11-4653/n.2019.10.014

文 / 姜涵予

纪录片的制作在新的时代要求能够结合时代发展的需要，充分利用目前快速发展的大数据技术，促进纪录片的形式表达与内容呈现。

1. 大数据技术

大数据技术主要包括数据挖掘、人工智能、自然语言处理、认知计算、机器学习、大规模数据处理等多种技术，使用过程中能够对大量的数据进行有效的清洗融合，并对大量数据进行建模分析，实现对大数据的建模与分析，建立大规模处理框架，并实现对其的可视化分析。在纪录片内容呈现与表现过程中这一技术始终贯穿其中，其中运用了多项分析技术，能够有效提升其分析能力[1]。

2. 大数据技术在记录品制作中的运用

2.1 总体技术运用

如此在记录片的内容分析与处理过程中，建立了基础的大数据存储库，包括 Cassandra/HBase 等 NoSQL 数据库、Hadoop 等分布式文件系统，在数据开发平台的运用过程中建立了以 Kafka 为代表的分析平台，在大数据处理框架的构建过程中则以 MapReduce，Spark/Storm 等作为代表，在这些方式的联合运用之下构建了高可扩展的、高可用性的处理架构，在开源大数据分布式并行处理分析过程中将其作为主流方式之一[2]。

在大数据分析过程中，在人工智能与机器学习充分发展的情况下，长短时记忆网络（LSTM）、卷积神经网络（CNN）、双向长短时记忆网络（BLSTM）、循环神经网络（RNN）等深度学习方式与记忆方式在纪录片的数据分析与存储过程中得到有效的开发与运用,实现了对视频、图像、文本、音频等多种模态数据分析的准确性与有效性，提升了纪录片展示出图像的丰富性与层次性[3]。

同时，在运用过程中实现了认知计算与知识图谱（Knowledge Graph）等层面的有效结合与运用，能够实现对亿级规模知识节点、大数据时代千万级及其相关关系之间的有效处理。在对大量复杂图像、技术的分析过程中运用了众包技术，同时在对各项信息、知识的分析过程中运用了多项数据技术方式，并且在图像与视频制作过程中能够进行数据的匿名制作，促进数据的安全性与隐私性。

2.2 查询引擎

Solr 技术，是在 Apache Lucene 基础之上开发的搜索方式，Instagram、Zappos、西尔斯、Netflix、Zappos、Travelocity、StubHub、eHarmony、百思买等用户也在运用这一搜索方式。

Presto，能够对超过 250PB 的数据进行交互式分析，属于一种开源的分布式 SQL 查询引擎，能够对纪录片中大量的图案信息进行交互式分析，具有较高的分析速度，具有比 MapReduce 与 Hive 更高的分析速度，分析性能较好。Drill 能够实现对 NoSQL、Hadoop 数据库进行有效查询并实现云储存。能够同时运行在上千个节点的服务器集群之中，在几秒之内快速处理 PB 级或者万亿条的数据记录。同时，能够实现对大量数据的挖掘和即时查询，谷歌云存储、MongoDB、MapR-FS、Azure Blob Storage、MapR-DB、Swift、HDFS 等多项数据库均可运用。

2.3 出具分析报道

Kibana 为 ElasticSearch 与 Logstash 日志分析中的 Web 接口，属于一种搜索仪表板与 Elasticsearch 分析方式，运用过程中能够对大量的信息进行高效搜索、可视化运行与有效操作[4]。

KNIME 即 Konstanz Information Miner，为开源分析和报表平台，属于一种康斯坦茨信息挖掘工具，能够同时实现对大量数据的有效处理，

Zeppelin 可运用于对 Python（Apache Spark）、Markdown、Apache Spark、Hive、SparkSQL 等格式的处理，是一种可进行交互数据分析的 Web 笔记本，在对数据库呈现与数据分析层面具有显著作用[5]。

Kettle 的运用能够对大量的图形信息等进行有效描述与分析，可用于多种不同数据库的数据管理，在 Pentaho 中具有重要作用，在目前国内多个项目中均有运用。见表 1 与表 2。

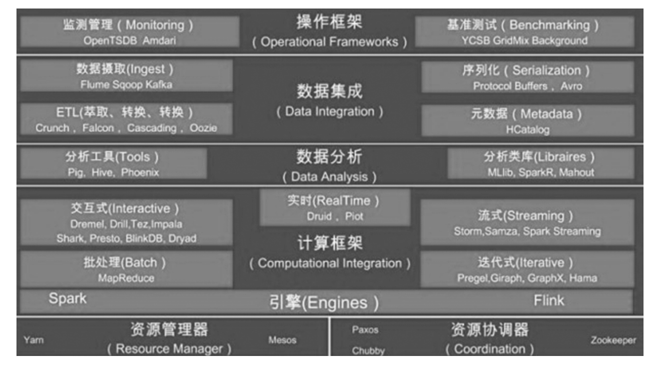


图 1　大数据架构图

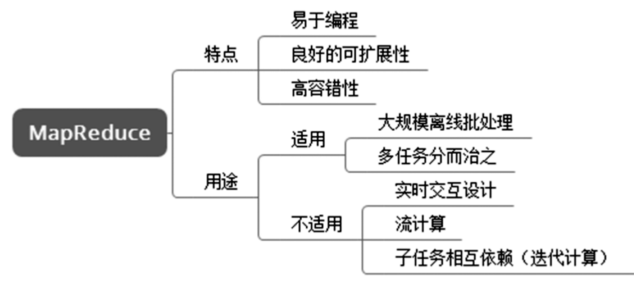


图 2　MapReduce 思维导图

3. 大数据技术驱动的纪录片内容呈现技巧变化

3.1 使用便携 4K 设备进行拍摄与数据处理

为了保证纪录片拍摄流程的完整性，可以运用便携 4K 设备进行拍摄，运用大数据技术与计算机技术对数据进行有效处理。

主机位选择一台 Sony F55，纪录时采用 4K XAVC 25P 格式，辅助机位选择另外一台 Sony FS7，纪录时采用 4K XAVC—l SOP 格式。在这两台拍摄机器之外，选择另外一台 Sony A7S2，录制时采用 4K XAVC 25P 格式，在机动机位的选择上设置的是同时具有良好灵活性并兼顾低照度环境的机位。如此就可以在拍摄过程中实现对常规环境、场景以及人文与不同人物之间对话的拍摄[6]。

除了主要的记录场景之外，纪录片拍摄过程中还需要对多种场景、整体环境、星空、延时摄影等环境进行拍摄，为此，在拍摄时设置两台尼康 D5200、一台 1DX、一台 SD2、两台 A7R2 等机器，同时搭配广角镜头与鱼眼进行拍摄。随着目前拍摄技术的发展，运用“精灵”等航拍器，可以对所需要拍摄的野外牧羊、古城遗迹、沙漠风光、村落生活等拍摄。在野外拍摄过程中，运用 4K 运动摄影机进行拍摄，例如 Go Pro 运动摄影机，运用车载、悬挂、手持等方式进行场景拍摄，可以解决拍摄过程中多活动场景机位不足以及室内空间较为狭窄等问题，同时在礼拜仪式、婚礼等场景的拍摄过程中也同样适用。在沙漠中一些小型动物的捕捉拍摄过程中也可以运用 GOPRO[7]（表 3）。

在纪录片《最后的沙漠守望者》拍摄过程中，除了前期摄影机参数解决方案之外，也为摄影机设置了图像指标，基本上是根据所选择的机型来确定。为此，《最后的沙漠守望者》拍摄过程中为 S—Gamut3.Cine 色域搭配 S—L093，采用 XAVC—S、XAVC—I 的视频压缩编码，两种方式在设置中的采样量化上有一定区别，但是不影响其后期调色，依然能够较为有效地匹配画面。这样结果的达成是由于牺牲了掉 RAW 格式纪录，能够在实现摄影组机动性的同时，实现最佳的 4K 组合。这一设备的运行能够实现 4K 兼顾高机动性，使摄影师与设备能够独立运行与拍摄，避免受到被拍摄对象的干扰。在调色方向上运用的 F55 与 FS7 的编码，在同时具有光比的市外环境下兼顾亮暗部细节。在较暗的环境下拍摄时，即使不打灯运用 A7S2 高 ISO 也能够获得较为清晰的画面，其中所运用到的设备有小型航拍器以及 GOPRO、F55、A7S2，在拍摄过程中摄影师均可单独作战，拍摄过程具有良好的流动性[8]。



图 3　纪录片《最后的沙漠守望者》

3.2 优化作业流程

在纪录片拍摄过程中流程具有重要作用，DIT 环节在前期摄影环节具有显著作用。在记录片的拍摄与制作过程中要求建立良好的 DIT 流程，导演、摄影以及后期导演等不同的职务人员能够各司其职，共同促进纪录片的拍摄。

为纪录片的拍摄制定严谨的 DIT 流程，要求拍摄素材具有良好的有序性与安全性，并将拍摄的作品第一时间反馈给摄制组，及时发现问题，并对此有效回馈，及时发现拍摄中存在的问题，并建立对其的改进方式，建立对其的有效保障机制，以此优化对其的后期处理，加强其前期技术与后期技术之间的有效串联与技术磨合。例如，在《最后的沙漠守望者》纪录片的拍摄过程中，DIT 流程的工作重点之一就是数据备份与整理。沙漠中的拍摄环境较为恶劣，可能出现较大的风尘阻断电力供应，对设备的良好运行会造成不良影响，影响了备份安全。为此，在纪录片的拍摄过程中采用了不需外接电源的大容量移动硬盘以及 MBP 笔记本进行备份，在数据备份上采用的是传统的一主一备方式，通过笔记本自身电池的供电能够实现硬盘正常的读写作业[9]。

将大数据的运用与设备的使用结合在一起，首先优化数据备份，通过后台程序的自动校验进行数据备份。验证备份的进行可以检查所需要备份的文件在物理结构上是否出现损毁，以此实现备份中的数据是否可读、可还原，以此实现备份文件的安全可靠性。

为了便于大数据的处理，制作素材管理查找表，将场记单中的有效素材进行充分筛选，结合拍摄需要进行

有效检查，并进行登记录入，使后期人员可以从中进行查阅。

完成素材收集之后，结合编辑制作的时间线，进行声话对位，结合天储存时间线，将 XML 表导出，使得在总部的剪辑制作过程中可以结合时间线进行编辑与制作，节省了大量素材过滤与整理的时间，提升了编辑制作效率[10]。

3.3 加强素材调色

在纪录片的拍摄过程中要结合导演的拍摄需求与纪录片的拍摄目的进行调色，以此来有效检查素材的拍摄质量，同时也能促进纪录片的制作组对纪录片的整体画面形成较为良好的认知风格，以此运用大数据等相关数据进行图像的多种表达，提升图像展示的层次性。

调色中的重要组成部分是预调色，以此来弥补前期拍摄中存在的一些瑕疵与不足，并且将其反馈给后期的摄影制作团队，结合预调色的拍摄呈现效果及时调整有效拍摄参数，并将其反馈给导演，帮助导演确定片子的风格。为此，在拍摄过程中采用 S—Gamut3．Cine 搭配 S—L093，这一技术的运用能够促进后期调色时还原人的肤色。见表 4。






图 4　预调色前后对比

在沙漠题材纪录片的制作过程中，由于沙漠中空气的通透性较为有限，在制作过程中针对一些较暗的部分调色难度较大，对此可以采用向右曝光的处理方式进行调色，保留较多的高光细节，在制作过程中尽量多保留一些较亮的区间进行图像的记录，较亮的部分具有更为优质的信噪比，能够达到最噪点的有效控制，在调色时能够达到较为良好的效果。

例如，在《最后的沙漠守望者》这一纪录片制作过程中采用了后期 DIT 数据格式交互的制作流程，实现了对大数据的有效利用。整体的纪录片类型包括 XAVC—S 带 LogMP4 文件、XAVC 带 LogMXF 文件等文件形式，视频种类多样，达到了 13T 左右，具有 100 多个小时的总体数量。文件类型同时包含采用延时摄影制作方式制作的 ProRes422HQ MOV 文件、H.264 压缩之后的 MOV 与 MP4 文件。在对这一纪录片制作与剪辑过程中，采用了 Premiere 进行 DIT 文件上线，支持多种剪辑方式，在导入之后即可进行剪辑，支持多种剪辑方式，在接导之后即可进行剪辑，省去大量的转码时间，提升了纪录片的剪辑效果。在 Link 模式运行过程中运用 Premiere 呈现，之后的运行过程中给达芬奇套底回批调色，以此促进制作、剪辑与制作。从前期拍摄一直到后期的制作、文件调整以及交互上线等均贯彻了 DIT 流程制作思想与制作方式。

结语

目前，大数据技术已经深入至生产、生活的各个领域之中，大数据技术为行业发展提供了新的技术、方式与支持。在纪录片的制作过程中运用大数据技术能够有效提升展示的层次性，促进纪录片形式的表达。在纪录片拍摄过程中要求制定严谨的 DIT 流程，要求拍摄素材具有良好的有序性与安全性，优化数据备份，通过后台程序的自动校验进行数据备份。完成素材收集之后，结合天储存时间线，将 XML 表导出，运用大数据等相关数据进行图像的多种表达，提升图像展示的层次性，优化纪录片的内容呈现。媒

参考文献

[1] 仝素梅 . 试谈”大数据”时代的计算机信息处理技术 [J]. 电脑编程技巧与维护，2017（22）：48-50.

[2] 张行健 . 基于数据库视角下解读大数据的研究进展与趋势 [J]. 电脑编程技巧与维护，2017（3）：43.

[3] 任宁 . 探讨基于教育大数据实现自主学习模式的挖掘技术 [J]. 电脑编程技巧与维护，2017（11）：51-53.

[4] 汤洪 . 关于大数据时代下网络新闻的写作技巧探讨 [J]. 传播力研究，2017（7）：138.

[5] 赵怀艳 . 纪录片“预售”浅析——做好“预售”的关键点 [C]. 第 35 届中国气象学会年会 S14 大数据、互联网、融媒体时代气象服务的创新与变革——第八届气象服务发展论坛，2018.

[6] 马伟平，李思琪 . 大数据时代传统媒体新闻报道的革新——以央视“据”说系列报道为例 [J]. 电视研究，2017(5)：65-66.

[7] 呼东燕 . 新时期央视文化类纪录片主题表达独特性分析——以《茶，一片树叶的故事》为例 [J]. 传播力研究，2017（9）：81-82.

[8] 王金环，李艳，腾明辉 . 基于大数据的地铁信息服务需求快速分类研究 [J]. 电脑编程技巧与维护，2018（3）.

[9] 高旗 . 如何利用大数据技术推进广播电视监测监管工作 [J]. 科技创新导报，2017（15）：180-181.

[10] 高昊，黄海蓉 . 浅析大数据时代纪录片内容生产的创新——以 NHK 纪录片《震灾大数据》为例 [J]. 中国电视，2018，387（5）：78-81.

（作者单位：吉林广播电视台）